

Marek Placiński

Nicolaus Copernicus University in Toruń
marpla@doktorant.umk.pl

Monika Boruta-Żywiczyńska

Centre for Language Evolution Studies, Nicolaus Copernicus University in Toruń
monika.feba.boruta@gmail.com

Investigating Natural Word Order via Pantomime: Research Report

Abstract. Inquiry into language evolution has recently focused on the question of the natural word order, i.e. a word order which may be primary in a cognitive and phylogenetic sense (Dryer, 2005; Pagel, 2009; Gell-Mann and Ruhlen, 2011). Some substantial insights into this topic originate in gesture and sign studies. Research by Goldin-Meadow et al. (2008) has inspired scientists to use the silent gesture paradigm, which requires participants to narrate events using their hands. The results of the study revealed that participants tended to produce SOV word order of a transitive event, regardless of the syntax of their native language. The finding was corroborated to a degree in later studies; however, some of them shed more light on the issue (Gibson et al., 2013; Hall et al., 2013; Sandler et al., 2005). The aim of our study is to test whether the SOV order is dominant when participants communicate transitive events (verbs) with whole-body pantomime.

Keywords: pantomime; natural word order; language evolution; linear grammar.

Natural Word Order

The question that we address in this paper is whether there exists any natural word order; that is, whether all humans, notwithstanding the language they speak, represent events in the same order: agent—patient—action, or

more formally, subject (S)—object (O)—verb (V). There is evidence from emerging sign languages (Senghas et al., 2004) and research on pantomime (Gibson et al., 2013; Hall et al., 2013; Sandler et al., 2005) which support this view. We base our assumptions on empirical experimentation conducted prior to our study.

All languages of the world are said to have developed from a single ancestor. Some evidence from comparative linguistics, such as the existence of cognates between different language families (Bengtson & Ruhlen, 1994, p. 281), seems to support this hypothesis. This ancestral language might have been spoken by the first behaviourally modern humans; alternatively, the single ancestry may be attributed to the fact that a bottleneck effect occurred later in human history, which means that one language was spoken at this particular period. The language is argued to have been a subject—object—verb language, which is supported by data from historical linguistics (Gell-Man & Ruhlen, 2011, p. 17290). For instance, Proto-Indo-European is claimed to have been an SOV language (Hock, 2015). Uralic languages, before the influence of the Indo-European languages, which over the course of history developed a SVO sentence structure, were also dominantly SOV. The same can be said about the languages from the Nostratic family and Afro-Asiatic families, and the remaining world language families (for a detailed discussion, c.f. Gell-Man & Ruhlen, 2011). The analysis of word order across language families leads Gell-Man and Ruhlen (2011) to suggest that the ancestral language had an SOV word order, which evolved into SVO order, out of which VSO and VOS emerged (Gell Man & Ruhlen, 2011, pp. 17292–17295).

Additionally, the topic of the natural word order is also crucial in language evolution research. Bickerton (1990), for instance, claims that proto-language did not have any rules which would govern ordering of elements and that language units were strung like beads on a string. Jackendoff and Wittenberg (2017) agree that linguistic communication at an early stage indeed did not have syntactic phrases or categories, and that it lacked such features as aspect or tense. However, they propound the view that such a linguistic communication had semantic rules for “mapping from semantic notions to linear order in phonology”, called linear grammar. Among the rules of such a grammar could state that the agent should always precede the patient (Jackendoff & Wittenberg, 2017, pp. 219–221). Taking these two arguments into consideration, our research addresses one of the central issues in language evolution, namely the existence of natural world order stemming from an earlier rudimentary form linguistic communication.

Other evidence suggesting that there is a natural word order comes from sign languages. Some of these languages, for instance, the Nicaraguan Sign Language (NSL) and Al-Sayyid Bedouin Sign Language (ABSL) have no syntactic model for order of words, yet they manifest a verb-final pattern. Data from both NSL and ABSL constitutes solid evidence in favour of the existence of a natural word order because these languages emerged without any external influence. NSL emerged among deaf children who previously had no contact with any standardised sign language and who were brought together to attend the same school, whereas ABSL among people who became deaf at a pre-linguistic stage (Sandler et al., 2005: 2664). Moreover, it has been found that children who are deprived of hearing and have no access to existing sign languages, tend to communicate using predicate-final pattern (Goldin-Meadow & Feldman, 1977).

In conclusion, the empirical study, comparative and historical reconstruction, and evidence from sign languages suggest that SOV is the default word order. However, although it is the most common one among languages of the world (Tomlin, 1986), it is not the only one. To answer the question why this is the case, new studies have been conducted.

The First Enquiry into the Natural Order of Events

The question Goldin-Meadow and colleagues (2008) asked in their research enquired about the influence language we speak has on our behaviour even when we are not speaking out. To answer it, they invited native speakers of four languages—English, Turkish, Spanish, and Chinese (Mandarin)—to perform two nonverbal tasks: a communicative task (describing an event by using gesture without speech) and a non-communicative task (reconstructing an event with pictures). The team assumed that if the language we use verbally has a structural pattern “assigned” to it, the realisation of the same message via nonverbal communication, will naturally reflect the pattern employed in the spoken language.

Testing the Hypothesis

The language groups were given two nonverbal tasks, based on previous findings of the researchers (Goldin-Meadow et al., 1996), the participants were asked to describe the events in speech before describing them in gesture, to determine their predominant speech order:

(1) **Gesture Task**, in which ten speakers of each language ($n=10 \times 4$) were asked to describe 36 pictures (simple motion events, in-place and crossing-space¹) displayed on a computer screen (depicting interactions between real objects and people and/or depicting animated toys) only by hand use. The task was designed to assess the ordering of words in nonverbal representations (Goldin-Meadow et al., 2008, p. 9163).

The focus of Goldin-Meadow et al.'s task was set on the position of semantic elements traditionally used to characterise word order in the world's languages: actors (Ar: Subjects, S), patients (P: Objects, O), actions (A: Verbs, V) (Gentner et al., 2001). The results showed that the speakers of the four languages consistently used ArA (SV) order when describing intransitive actions, but different orders to describe transitive events. The participants employed the orders of their languages, English and Spanish speakers used ArAP (SVO) and Turkish speakers used ArPA (SOV) to describe all transitive actions. Chinese group employed ArAP (SVO) for in-place events but ArPA (SOV) for crossing-space transitive events (Goldin-Meadow et al., 2008, p. 9163).

The results, presented for in-place and crossing-space events with one within-subjects factor (order) and one between-subjects factor (language group), showed that the effects were significant for order but not group in each analysis: gesture strings were significantly more likely to display the ArPA (SOV) order than the ArAP (SVO) order for spoken Chinese, English, and Spanish (in-place actions) and in spoken English and Spanish (crossing-space actions). Further, the researchers found that modality influenced ordering – gesture was significantly different from speech for English, Spanish, and Chinese, but not for Turkish. The participants did not display the order found in their spoken language in their gestures. Instead, the gestures all followed the same ArPA (SOV) pattern (Goldin-Meadow et al., 2008, pp. 9163–9164).

(2) **Transparency Task**, in which another ten speakers of each language ($n=10 \times 4$) were asked to reconstruct the same events by using sets of transparencies, each presenting a different object or action. The participants were asked to stack images on one another to reconstruct the event. The task was designed to test if speakers would extend the ordering patterns of their languages not only to the pictorial modality (Goldin-Meadow et al., 2008, p. 9163).

¹ Examples of sentences devised by Goldin-Meadow et al. (2008) include: crossing space: duck-moves-to-wheelbarrow, girl-gives-flower-to-man; in-place: woman-twists-knob, man-plays-guitar; (full list of examples, see: <http://www.pnas.org/content/suppl/2008/07/01/0710060105.DCSupplemental.DOA:12.11.2018>).

In their answers, the participants consistently employed an order of stacking transparencies, and the structure remained the same across languages. Intransitive and transitive events were analysed separately using one within-subjects factor (order) and one between-subjects factor (language group) ANOVA. As a result, the ArA (SV) order was present much more often than any other structure for both kinds of intransitive events, and ArPA (SOV) for both kinds of transitive actions. Importantly, speakers of the four language did not betray the syntactic structures of their spoken languages in their nonverbal re-production, and consistently employed the ArPA (SOV) pattern (Goldin-Meadow et al., 2008, pp. 9163–9165).

The researchers found that the word orders speakers used in their everyday speech did not influence their nonverbal behaviour. Interestingly, the speakers of all four languages employed the same word order in both nonverbal tasks. The authors claim that the actor–patient–action (subject–object–verb) order is found in many languages of the world as well as in newly developing gestural languages. Their results give insight into the natural order of language – the one that seems universal across languages of the world while re-producing events anew (Goldin-Meadow et al., 2008, pp. 9164–9166).

New Evidence

After Goldin-Meadow et al. (2008), more experiments have been conducted, for instance Gibson et al. (2013) and Hall et al. (2013). These two studies used a similar experimental paradigm, however, the scenes that participants were supposed to re-enact involved either one animate object (“the man”), one inanimate object (“the ball”), and an action (“kicks”), or two animate objects (“the man”, “the dog”) and an action (“bites”). The first type of events are called “non-reversible”, since only one of the objects can perform the action in the scene (as only the man can kick the ball, not the other way round). The second type, on the other hand, is called “reversible” events because both objects involved in the event can perform the action (as both the dog and the man can bite). As a consequence of changing the experimental paradigm, new results were found, which may explain why, despite SOV being the “default” world order, SVO is so prevalent.

The Noisy Channel Hypothesis

The first study which involved non-reversible and reversible events, and which aimed to address the issue of SVO prevalence was conducted by Gibson et al. (2013). They proposed two reasons why SVO might be favoured over SOV. The first reason why SOV becomes dispreferred is because of memory constraints. Secondly, Gibson and colleagues hypothesise that language production and comprehension operate in a noisy channel (Shannon, 1948); i.e. a message is corrupted when it is sent via the channel. Therefore, in order to ensure that the recipient will decode the intended message with minimal effort, the sender chooses the best strategy which will allow the recipient to decode the message (Gibson et al., 2013, p. 1080).

The use of reversible and non-reversible events to test the noisy-channel hypothesis in the experiment was to ensure that there are messages which are easy to decode notwithstanding the order of elements (non-reversible events) and more complex ones, in which noise may prevent a straightforward interpretation of messages (reversible events). In the first case, the noun phrase representing the doer of the action cannot be confused with the object; however, in the latter case either noun phrase can be interpreted to be the subject, since both *the man* and *the dog* can *bite* someone. Hence, with non-reversible events, the sequence of communicating information is not that important. However, it may play a role in reversible events, and SVO may be the preferred way to communicate. Employing SVO has another positive value for comprehension because deletion only partially obscures a message and recovering the meaning is not as difficult in comparison with using SOV for reversible events. If the object is lost, the recipient still understands that *the man bites*; equally, when the subject is lost, the recipient also knows that someone *bites the dog* (Gibson et al., 2013, p. 1080).

Gibson and colleagues invited participants who were native speakers of English (SVO), Korean (SOV), and Japanese (SOV). The participants were shown silent animations of transitive events, both reversible and non-reversible ones. First, they had to verbally describe the situation, and later on gesture it as well (Gibson et al., 2013, pp. 1080–1081).

The results Gibson et al. found indicate that when the participants had to gesture non-reversible events, they preferred SOV word order. However, when they had to do the same with reversible events, this preference disappeared in favour of SVO word order. The data obtained in the experiment lead Gibson and colleagues to suggest that senders communicate reversible events using SVO order to ensure that the message is easy to be decoded.

Some participants who nevertheless gestured reversible events in an SOV manner developed strategies which resembled case marking in natural languages (such as gesticulating “1” for the subject and “2” for the object). Moreover, Gibson et al. apply the same reasoning for the presence of SOV and SVO languages. Because of the noise in the channel, which obscures decoding a message, non-SVO languages have developed mechanisms (such as case-marking), which make the preservation of SOV word order possible. However, if no such strategies are developed, languages shift towards SVO syntax (Gibson et al. 2013, p. 1085).

Cognitive Constraints

A study conducted by Hall and colleagues (2013), which they conducted after Gibson et al.’s (2013), also aimed at explaining why SOV word order is not suitable for representing reversible events in communication, despite being the default order in which the human cognitive system orders elements. Hall et al. decided to address the question of natural word order because they did not agree with the conclusions that Gibson et al. (2013, described above) and Meir et al. (2010) arrived at, and because of the discrepancies between conclusions of these studies.

Meir et al.’s (2010) study examined two groups, one being a group of native speakers of Hebrew (SVO) and the second of Turkish (SOV). Participants of the experiment were supposed to communicate, using pantomime, reversible and non-reversible events. Similarly to Gibson et al.’s (2013) findings, Meir and colleagues found that, although SOV is the preferred word order when in the case of reversible events, participants tended to shift towards SVO representation. They situated their study in the area of functionalist linguistics and attributed the shift to avoiding confusability, since they assumed that the roles of the agent and the patient in *the man the dog bites* are more readily confused than in the case of *the man bites the dog*. Meir et al.’s assumptions can be said to strictly concern comprehension processes. In here, when the sender sees that the recipient has problems with decoding the intended message, he has to invent a new strategy to communicate more efficiently (Hall et al., 2013, pp. 3–4).

The discrepancies between the two studies stem from the way participants represented events and from the theoretical findings in the studies. In Meir et al.’s study, some participants used OSV in pantomime, which goes against the confusability hypothesis, because two animate constituents are near one another. In contrast, such cases were not as common in Gibson et al.’s

experiment; nevertheless, it is equally difficult to explain OSV under the noisy channel hypothesis. Therefore, to address the theoretical and empirical discrepancies between the two studies, Hall and colleagues conducted another study (Hall et al., 2013, p. 4).

In the study, Hall et al. invited native speakers of English to describe events in pantomime in two conditions. In condition one, reversible and non-reversible events were mixed. In condition two, participants had to first describe non-reversible events, and after that reversible ones (Hall et al., 2013, p. 4).

The results Hall and colleagues obtained in the experiment were similar to the ones found in previous studies: although SOV word order was common for non-reversible events, participants tended to shift to SVO when they had to re-enact reversible events (Hall et al., 2013, p. 13).

On a more theoretical plane, the study challenged the two previous hypotheses concerning a shift towards SVO, and it proposed a new explanation. The reservations the authors had to the confusability hypothesis (Meir et al., 2010) was that it has not been tested whether it really is more difficult to process reversible events which are re-enacted using the SOV word order. Concerning the noisy channel hypothesis, Hall and colleagues argue that it is incompatible with current findings in research on language production. Currently, production vulnerabilities are believed to have influence on language production, and not, as the noisy channel hypothesis assumes (Gibson et al., 2013), that comprehension vulnerabilities influence producing language. Because of these inconsistencies, Hall et al. proposed a new hypothesis: role conflict in production (Hall et al., 2013, p. 13).

A role conflict occurs when a person re-enacts reversible events in pantomime. In such re-enactments, participants typically use their bodies to perform both the role of the agent and the patient. Here, gestures which correspond to actions force them to embody the role of the agent. Situations in which patient is re-enacted right after the agent are problematic because of shifting roles; in other cases, such as SVO, OSV, SOSV, and SOSVO, the problem disappears. Thus, the choice of avoiding SOV word order does not stem from its potential ambiguity, but rather because producers recognise that there might be a mismatch between the role of the agent and the role they have recently adopted, and this is why they decide to produce the action gesture right after the agent (Hall et al., 2013, p. 13).

Hypothesis

Our hypothesis assumes that one of the patterns will dominate in the pantomimic re-enactments of the scenes that participants were supposed to re-enact. Based on the abovementioned research of Goldin-Meadow and colleagues, we think SOV may turn out to be the strategy unconsciously chosen by the actors as a more communicative one. On the other hand, the pattern that emerged in Hall et al.'s and Gibson et al.'s research contradicts Goldin-Meadow et al.'s results and suggests that the pattern of spoken language is present in pantomimes and hence facilitates understanding of the message conveyed via body movements.

Methods

Twelve adults (students-actors; aged 18–21, both men and women (M=5, W=7), native speakers of Polish language, took part in the experiment. They were each instructed to pantomimically re-enact transitive events (n=20) they were presented with on a piece of paper, as indicated in Hall et al.'s (2013) and Gibson et al.'s (2013) research. The material was prepared originally in the form of hand-drawn cartoons (see below, Fig. 1), based on our previous research on syntactic strategies developed in pantomime (Boruta & Placiński, 2017). The material, a printed matrix, displayed 10 types of simple events—*kick, wave, throw, walk, look, hold, shout, push, twist, pull*—in two conditions: a reversible one (n=10; the agent and the patient are animate) and a non-reversible one (n=10; the agent is animate, and the patient is not). The actors were supposed to present only one event at a time and familiarise with a new picture upon completion of the former one. They were asked to look at the picture (for 20 seconds) and re-enact the scene (20 seconds) in front of a camera (Sony HDR-CX405B) in one of the rooms in Collegium Humanisticum of Nicolaus Copernicus University in Toruń, Poland. Although Goldin-Meadow et al. (2008) specify that spoken language can influence gesture (or, in our case, pantomime), we did not ask the participants to verbalise what they saw in a picture. This decision was made based on the studies discussed earlier in the paper—we aimed at avoiding a situation in which verbalised content of the picture would interfere with nonverbal realisation of it.

We established a time frame for preparation and presentation (20 seconds each), as we wanted to remain a steady level of the preparation-presentation

time between the participants, and prevent some of the actors from taking too much time devising their strategy. We intended that, if a strategy shall arise, it will appear in a more natural or improvised manner.

In total, we obtained 240 short videos of equal number of reversible and non-reversible events ($n=120$ for each condition). The videos were further analysed in ELAN (4.8.1 version) annotation tool by two expert judges who assessed the word order for each scene and later checked the congruence of their choices. We did not employ Cohen's kappa coefficient, as timing of actions was insignificant for the assessment. The results, grouped into syntactic strategies, were checked for significance in SPSS Statistics. We used t-test for independent samples to check whether any of the strategies was prevalent.



Figure 1. Stimulus: sample events in the form of hand-drawn cartoons

Results

Our results contradict the previous findings due to the fact that there are no major differences between the syntax of pantomime and the syntax of spoken language. Such a discrepancy may result from the nature of the Polish syntax which has a relatively free word order and which often drops the subject pronoun due to verbal inflections (Embick, 1995, Szczegielniak, 2001). The participants also performed re-enactments which did not conform to the standard SVO, OSV, SV word orders which held the majority. The number of outlier cases for reversible events amounted to $n = 7$ (cases: SOSVO, only S / only O), and for non-reversible events to $n = 6$ (cases: OOSV, OSOSV, O). Interestingly, we found cases of gender and case marking. For instance, two male participants indicated that they were acting a female role by gesturing long hair. Case marking was also attested – two actors, one male and one female, indicated that they were performing the subject or the object role signalling that by the use of erected index finger (first person) or two erected fingers (second person). In some cases, the two strategies were combined by the actors.

We observed no differences between groups in terms of significance – none of the patterns was prevalent. We expected a tendency towards one or the other strategy, however, t-test result for independent samples was: $t(240) = 0.635$; $p = 0.526$ and proved to be insignificant.

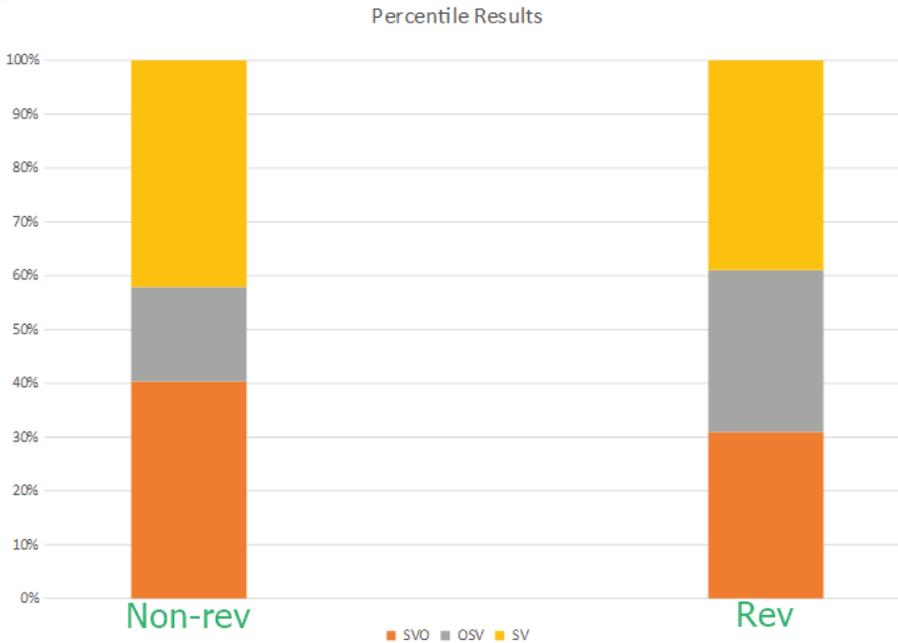


Figure 2. Results – percentage of patterns (SVO – Orange, OSV – Grey, SV – Yellow)

Such a result may stem from several factors:

1. The syntactic pattern of the Polish language (the dominance of the SVO pattern), but/and its fairly liberal structure (Embick, 1995, Gumul, 2011);
2. The lack of the need to make the scene as communicative as possible (consider a game of charades when participants have to transfer meanings efficiently due to the limited time);
3. The focus on such aspects as gender or case marking, which occurred consistently in some actors' re-enactments.

Interestingly, if we add up SVO and SV patterns, they hold the majority in reversible and non-reversible events, which may suggest that the pattern that is present in spoken language is reflected in pantomimic re-enactments of events. To confirm that idea, however, we would have to collect data on spoken realisation of each scene from each actor. On the other hand, as

Goldin-Meadow et al. (2008) note, such a procedure could force the actors into incorporating spoken language pattern in their pantomime.

Conclusions

The results we obtained in the experiment corroborate the role conflict hypothesis, since both SVO and OSV structures were attested in the re-enactments. Such results can be reconciled neither with the noisy channel nor confusability hypotheses because both of them assume that OSV should be avoided by participants in reversible events. Such a result conforms to the hypothesis that proto-language did not have syntactic rules – as it is troublesome to find fixed patterns in pantomimic narration of events. However, even rudimentary form of visual communication devoid of grammar fulfils the requirements needed for the purpose of message transfer. Further studies could employ research on communicativeness of the events re-presented pantomimically. We could look at the strategies actors use in their re-enactments – whether we can observe a learning effect and hence change. Additionally, we could assess the degree to which a consistent strategy facilitates the understanding of a message.

References

- Bengtson, J. D., & Ruhlen, M. (1994). Global etymologies. In M. Ruhlen (Ed.), *On the origin of languages: Studies in linguistic taxonomy* (pp. 277–336). Stanford: Stanford University Press.
- Bickerton, D. (1990). *Language & species*. Chicago-London: University of Chicago Press.
- Boruta, M., & Placiński, M. (2017). The syntax in pantomimic re-enactments of events among Polish participants. *Culture and Education Journal*, 2(116), 106–118.
- Dryer, M. S. (2005). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, & D. Gil (Eds.), *The world atlas of language structures* (pp. 330–333). Oxford: Oxford University Press.
- Embick, D. (1995). Mobile inflections in Polish. In J. N. Beckman (Ed.), *Proceedings of NELS*, 25(2), 127–142.
- Gell-Mann, M., & Ruhlen, M. (2011). The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, 108(42), 17290–17295.
- Gentner, D., Boroditsky, L., Bowerman, M., & Levinson, S. C. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.),

- Language acquisition and conceptual development, individuation, relativity and early word learning* (pp. 215–256). New York: Cambridge University Press.
- Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. *Psychological Science*, 24(7), 1079–1088.
- Goldin-Meadow, S., & Feldman, H. (1977). The development of language-like communication without a language model. *Science*, 197(4301), 401–403.
- Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. *Psychological Review*, 103(1), 34–55.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 105(27), 9163–9168.
- Gumul, E. (2011). Linearity constraint in simultaneous interpreting. *Linguistica Silesiana*, 32, 163–178.
- Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2013). Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition*, 129(1), 1–17.
- Hock, H. H. (2015). Proto-Indo-European verb-finality: Reconstruction, typology, validation. In L. Kulikov & N. Lavidas (Eds.), *Proto-Indo-European syntax and its development* (pp. 51–78). Amsterdam: John Benjamins Publishing Company.
- Jackendoff, R., & Wittenberg, E. (2017). Linear grammar as a possible stepping-stone in the evolution of language. *Psychonomic Bulletin & Review*, 24(1), 219–224.
- Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. A. (2010). The interaction of animacy and word order in human languages: A study of strategies in a novel communication task. *Proceedings of the Eighth Evolution of Language Conference*, 455–456.
- Pagel, M. (2009). Human language as a culturally transmitted replicator. *Nature Reviews. Genetics*, 10(6), 405.
- Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, 102(7), 2661–2665.
- Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, 305(5691), 1779–1782. doi: 10.1126/science.1100199
- Szczegielniak, A. (2001). Polish optional movement. In G. Alexandrova & O. Arnaudova (Eds.), *The minimalist parameter* (pp. 125–148). Amsterdam: John Benjamins Publishing Company.
- Tomlin, R. (1986). *Basic word order: Functional principles*. London: Croom Helm.